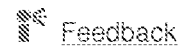



[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
Search: ☒ The ACM Digital Library ☐ The Guide


THE ACM DIGITAL LIBRARY


Recovery in the Calypso file system

Full text [Pdf \(319 KB\)](#)
Source **ACM Transactions on Computer Systems (TOCS)** [archive](#)

 Volume 14 , Issue 3 (August 1996) [table of contents](#)

Pages: 287 - 310

Year of Publication: 1996

ISSN:0734-2071

Authors [Murthy Devarakonda](#) IBM T. J. Watson Research Center, Yorktown Heights, NY
[Bill Kish](#) IBM T. J. Watson Research Center, Yorktown Heights, NY
[Ajay Mohindra](#) IBM T. J. Watson Research Center, Yorktown Heights, NY

Publisher [ACM](#) New York, NY, USA

Bibliometrics Downloads (6 Weeks): 2, Downloads (12 Months): 44, Citation Count: 5

Additional Information: [abstract](#) [references](#) [cited by](#) [index terms](#) [review](#) [collaborative colleagues](#) [peer to peer](#)
Tools and Actions: [Request Permissions](#) [Review this Article](#)
[Save this Article to a Binder](#) [Display Formats: BibTeX](#) [EndNote](#) [ACM Ref](#)
DOI Bookmark: Use this link to bookmark this Article: <http://doi.acm.org/10.1145/233557.233560>
[What is a DOI?](#)

↑ ABSTRACT

This article presents the design and implementation of the recovery scheme in Calypso. Calypso is a cluster-optimized, distributed file system for UNIX clusters. As in Sprite and AFS, Calypso servers are stateful and scale well to a large number of clients. The recovery scheme in Calypso is nondisruptive, meaning that open files remain open, client modified data are saved, and in-flight operations are properly handled across server recover. The scheme uses distributed state amongst the clients to reconstruct the server state on a backup node if disks are multiparted or on the rebooted server node. It guarantees data consistency during recovery and provides congestion control. Measurements show that the state reconstruction can be quite fast: for example, in a 32-node cluster, when an average node contains state for about 420 files, the reconstruction time is about 3.3 seconds. However, the time to update a file system after a failure can be a major factor in the overall recovery time, even when using journaling techniques.

↑ REFERENCES

Note: OCR errors may be found in this Reference List extracted from the full text article. ACM has opted to expose the complete List rather than only correct and linked references.

- 1 [O. Babaoglu, R. Davoli, L. A. Giachini, M. Gray Baker, RELACS: A communications infrastructure for constructing reliable applications in large-scale distributed systems.](#)